1. General Course Information

Course Information
Statistical Science 4850G/9850B. Advanced Data Analysis: Statistical Learning and Predictive Analytics.
Times & Location: MWF Rm 248, 12:30 to 1:30
Start Date: Wednesday January 9, 2019.

Prerequisite Requirements
A minimum mark of 60% in both Statistical Sciences 3843A/B and Statistical Sciences 3859A/B.

Unless you have either the requisites for this course or written special permission from your Dean to enroll in it, you may be removed from this course and it will be deleted from your record. This decision may not be appealed. You will receive no adjustment to your fees in the event that you are dropped from a course for failing to have the necessary prerequisites.

A knowledge of R/RStudio including the use of markdown to produce beautiful PDF technical reports is assumed or that you are willing to learn.

2. Instructor Information
A. I. McLeod, Ph.D., Professor
Office: WSC 235
Office: Hours MWF 2:30-3:30 or by appointment
Email: aimcleod@uwo.ca

Students must use their Western (@uwo.ca) email addresses when contacting their instructors. I can only respond to email about administrative matters dealing with the course. All questions about the course material must be discussed in person with me, either at the end of the lecture, during the lecture if it is a substantial point that is worth a full discussion in class (you will receive extra bonus credit for raising such questions) or during office hours.

3. Course Description/Syllabus
I will discuss many methods described in the textbooks (ESL, ISLR, APM) listed in the next section. I will also provide extensive class lecture notes. These notes will include R scripts and notebooks illustrating all the main methods discussed in class.

4. Course Materials
No course textbooks have been assigned but most of my lectures are derived from the three textbooks listed below. The ESL and ISLR books have freely available PDF copies available on the authors' website. Students should check OWL (http://owl.uwo.ca) on a regular basis for news and updates. This is the primary method by which information will be disseminated to all students in the class. Students are responsible for checking OWL on a regular basis.

ESL

http://statweb.stanford.edu/~tibs/ElemStatLearn/
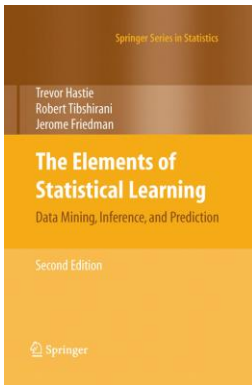
This is a celebrated book in the field of statistical/machine learning communities. It is covers a large number of topics with substantial discussion of the mathematical aspects to the methods and algorithms. The authors and their colleagues, especially Leo Breiman, have made substantial contributions to ML. PDF freely available.
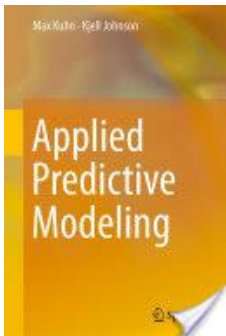
ISLR

http://www-bcf.usc.edu/~gareth/ISL/
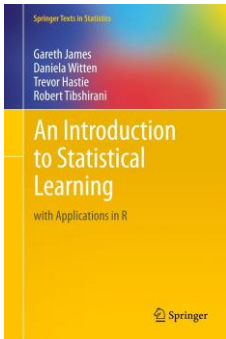
This is an easy to read simple account of a few of the most important methods discussed in the ESL book. A CRAN package ISLR is provided. PDF freely available.

APM

Covers many methods in ELS, more than ISL and with more depth and insight. Datasets are available in the CRAN package AppliedPredictiveModeling. The CRAN package caret implements many of the techniques from this book and an online tutorial provides a comprehensive overview: http://topepo.github.io/caret/index.html

5. Methods of Evaluation

The overall course grade will be calculated as listed below:
Participation   5%
MOOC   5%
Nonlinear Project   15%
Final Project   75%

Participation. Regular attendance is expected and attendance may be taken. This will be included in the participation grade. Contributing to the course by raising questions in class or with discussions with me after class will also be counted towards participation. Students who regularly do not attend lectures will receive a warning and if they persist will receive a failing grade.

MOOC. Completing this online courses by achieving 90% on the quizzes will validate your basic competence. If successful, upload your Certificate of Accomplishment with Distinction, to your OWL webpage to earn 5%. Both the MOOC and the certificate are freely available. To enroll see:
https://online.stanford.edu/course/statistical-learning-self-paced

Nonlinear Project (due Monday February 18, 2019). Involves at least one of the following prediction methods: k-nearest neighbour, loess, spline regression, logistic regression or MARS. Students are expected to find their own data. A beautiful PDF report prepared using RStudio markdown with latex or Word is required since the concept of "reproducible research" is important.

Final Project (due April 28 at the very latest). Involves applying the methods discussed in this course to some interesting data acquired by the student and producing a beautiful PDF report.

Note: all reports must be uploaded to the students' dropbox on OWL. Reports submitted by email cannot be accepted. Failure to reference the data source or sources or other relevant work may be regarded as a type of plagiarism and will be penalized. I will use suitable resources to ensure the quality and originality of the reports.

6. Accommodation and Accessibility

If you are unable to meet a course requirement due to illness or other serious circumstances, you must provide valid medical or supporting documentation to the Academic Counselling Office of your home faculty as soon as possible. If you are a Science student, the Academic Counselling Office of the Faculty of Science is located in WSC 140, and can be contacted at scibmsac@uwo.ca.

For further information, please consult the university's medical illness policy at http://www.uwo.ca/univsec/pdf/academic_policies/appeals/accommodation_medical.pdf.

7. Academic Policies
The website for Registrarial Services is http://www.registrar.uwo.ca.

In accordance with policy, http://www.uwo.ca/its/identity/activatenonstudent.html, a centrally administered e-mail account provided to students will be considered the individual's official university e-mail address. It is the responsibility of the account holder to ensure that e-mail received from the University at his/her official university address is attended to in a timely manner.

Scholastic offences are taken seriously and students are directed to read the appropriate policy, specifically, the definition of what constitutes a Scholastic Offence, at this website: http://www.uwo.ca/univsec/pdf/academic_policies/appeals/scholastic_discipline_undergrad.pdf.

8. Support Services
The policy on Accommodation for Students with Disabilities can be found here: www.uwo.ca/univsec/pdf/academic_policies/appeals/accommodation_disabilities.pdf

The policy on Accommodation for Religious Holidays can be found here: http://www.uwo.ca/univsec/pdf/academic_policies/appeals/accommodation_religious.pdf

Students who are in emotional/mental distress should refer to Mental Health@Western (http://www.health.uwo.ca/mental_health) for a complete list of options about how to obtain help. Additional student-run support services are offered by the USC, http://westernusc.ca/services.

This course is supported by the Science Student Donation Fund. If you are a BSc or BMSc student registered in the Faculty of Science or Schulich School of Medicine and Dentistry, you pay the Science Student Donation Fee. This fee contributes to the Science Student Donation Fund, which is administered by the Science Students' Council (SSC). One or more grants from the Fund have allowed for the purchase of equipment integral to teaching this course. You may opt out of the Fee by the end of September of each academic year by completing paperwork in the Faculty of Science's Academic Counselling Office. For further information on the process of awarding grants from the Fund or how these grants have benefitted undergraduate education in this course, consult the chair of the department or email the Science Students' Council at ssc@uwo.ca.